Purification of large bicolorable graph states

Kovid Goyal,[1,*] Alex McCauley,[1,†] and Robert Raussendorf[2,‡]

[1] *Institute for Quantum Information, California Institute of Technology, Pasadena, California 91125, USA*
[2] *Perimeter Institute, 31 Caroline Street North, Waterloo, Canada, N2L 2Y5*

(Dated: May 26, 2006)

We describe novel purification protocols for bicolorable graph states. The protocols scale efficiently for large graph states. We introduce a method of analysis that allows us to derive simple recursion relations characterizing their behavior as well as analytical expressions for their thresholds and fixed-point behavior. We introduce two purification protocols with high threshold. They can, for graph degree 4, tolerate 1% (3%) gate error or 20% (30%) local error.

I. INTRODUCTION

The known protocols in quantum information processing require a certain degree of quantum-mechanical entanglement to achieve an advantage over their classical counterparts. Often, this quantum-mechanical "essence" is provided in terms of in-advance-prepared quantum states. For example, Bell states are used in a well-known protocol for quantum cryptography [1], and schemes for multiparty cryptographic tasks using Greenberger-Horne-Zeilinger (GHZ) states and other Calderbank-Shor-Steane (CSS) states have been devised [2]. Further, in quantum computation, multiparticle entangled states can be used to streamline the execution of gates and subcircuits via gate teleportation [3], and cluster states represent a universal resource for quantum computation by local measurements [4].

In most realistic scenarios the quality of entangled resource states is degraded by the effects of decoherence and methods of error detection or correction are required to counteract this process. One such method is state purification where a (close to) perfect copy of a quantum state is distilled out of many imperfect ones. Purification was first described for Bell states [5, 6, 7] and subsequently generalized to bicolorable graph states and CSS states [8, 9, 10]. Recently, a protocol for the purification of W states was presented in [11]. State purification is used, for example, to establish a perfect quantum channel between two parties [5], to efficiently create long-range entanglement via quantum repeaters [12] or to render certain schemes for topological fault-tolerant quantum computation universal [13].

Imperfect initial states are not the only sources of error for realistic state purification. With the exception of certain schemes of topological quantum computation such as [13], errors in the gates for purification also need to be taken into account.

What can we expect to gain from an imperfect purification procedure? In the process of purification the errors of the initial state are replaced by the errors of the purifying gates. Thus, the amount of error may be reduced if the quality of the initial states is low compared to the quality of the gates for purification (but above threshold). Further, purification can be used to *condition* the error of a quantum state. For example, imperfect Bell-state purification can be used to establish a perfectly private if imperfect quantum channel [14]. In a multiparty scenario, for some protocols the purification gates act locally on each copy of the state to purify, resulting in a local or close to local error model for the final state. This feature attains relevance in the context of fault-tolerant quantum computation. Threshold theorems have been established for increasingly general types of error including coherent and long-range errors [15, 16] but there are realistic scenarios in which standard error correction appears to fail [17]. In such a situation, state purification may be used to turn the error model into a more benign one.

The focus of this paper is purification of bicolorable graph states by imperfect means, a subject that has previously been studied in [9, 18, 19]. We are interested in the interplay between threshold and overhead. Specifically, we seek protocols that, (I) work with erroneous purification gates, (II) have a high threshold and good quality of the output state, (III) scale efficiently, and (IV) are analytically tractable.

Hashing [2, 6, 10] protocols have a high threshold in the error of the initial state and require only a minimal resource overhead, but they break down as soon as the purification gates become slightly imperfect [26]. Recursive protocols such as [8] also have a high threshold for error in the initial states and furthermore work with imperfect purification gates, but they are exponentially inefficient in the number of particles.

Our protocols are resistant to initial as well as purification errors and are computationally efficient. As a bonus, our protocols are analytically tractable for a wide class of errors. Specifically, our base protocol described in Sec. III can be analyzed for arbitrary input states and general probabilistic Pauli errors in the purification gates. This fact arises through a special locality property. So far, the exponential increase of parameters in the description of n-particle mixed states—even mixed stabilizer states—has been found to be an obstacle to analytic discussion, and only severely restricted error

*kovid@theory.caltech.edu
†mccaule@caltech.edu
‡rraussendorf@perimeterinstitute.ca

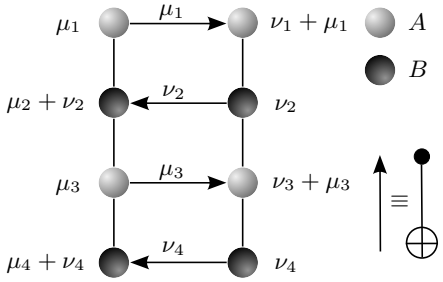

FIG. 1: Action of MCNOT in the graph basis. The arrows represent the direction of syndrome (or Z error) flow (i.e. the action of the MCNOT on the stabilizer)

models have been treated in the literature.

This paper is organized as follows: in Sec. II we briefly review the protocol [8] for purification of bicolorable graph states. In Secs. III and IV C we describe our purification protocols and characterize them in terms of purification threshold, output quality, and overhead. We conclude with a discussion of our results in Section V.

II. BRIEF REVIEW

Consider a graph $G(V, E)$ with vertex set V and edge set E. $G(V, E)$ is bicolorable if V can be partitioned into two disjoint subsets A and B such that every edge in E connects a vertex in A with a vertex in B. E defines a *neighborhood* relation on elements of V; $\mathcal{N}(j) := \{i \in V : (i\ j) \in E\}$. Define the correlation operators

$$K_j := X_j \prod_{i \in \mathcal{N}(j)} Z_i \tag{1}$$

where X, Y, and Z are the Pauli matrices. A graph state is a $|V|$-qubit state $|\boldsymbol{\mu}\rangle$ ($\boldsymbol{\mu} \in \{0,1\}^{|V|}$) that satisfies the eigenvalue equations

$$K_j |\boldsymbol{\mu}\rangle = (-1)^{\mu_j} |\boldsymbol{\mu}\rangle, \forall j = 1, \ldots, |V|. \tag{2}$$

The states $\{|\boldsymbol{\mu}\rangle\}$ form a basis of the Hilbert space of $|V|$-qubit states called the *graph basis*.

We now briefly discuss the post-selection protocol of [8]. The protocol works by taking two identical copies of a bicolorable graph state and performing multiple CNOTs (MCNOT) between them, in a definite pattern as illustrated in Fig. 1. Relabeling states in the graph basis to reflect the partition into colors A and B (i.e., $|\boldsymbol{\mu}\rangle \equiv |\boldsymbol{\mu}_A, \boldsymbol{\mu}_B\rangle$), the effect of the MCNOT is [8]

$$|\boldsymbol{\mu}_A, \boldsymbol{\mu}_B\rangle |\boldsymbol{\nu}_A, \boldsymbol{\nu}_B\rangle \mapsto |\boldsymbol{\mu}_A, \boldsymbol{\mu}_B + \boldsymbol{\nu}_B\rangle |\boldsymbol{\nu}_A + \boldsymbol{\mu}_A, \boldsymbol{\nu}_B\rangle, \tag{3}$$

where $+$ is elementwise addition modulo 2. Notice that information about $\boldsymbol{\mu}_A$ has been copied into state 2 and information about $\boldsymbol{\nu}_B$ has been copied into state 1. We then measure the local observables X and Z on copy 2, and reconstruct from the measurement outcomes the eigenvalues of all K_j with $j \in A$. Suppose we get -1 at the kth qubit. Then we know that either μ_k or ν_k was 1, but we do not have enough information to decide which one, so we throw away the states and start again. We keep doing this until all measurements are clear. By this procedure we correct, to lowest order, errors in the qubits of color A. In the next round we interchange the roles of colors A and B and so purify the B qubits. We can concatenate this procedure to achieve desired levels of purity. Because we are post-selecting states on the basis of a global measurement outcome, this protocol is inefficient for large states. This inefficiency can be addressed by using error correction instead of post-selection, to which we now turn.

III. THREE-COPY PROTOCOL

The simplest way to get enough information to perform error correction is to do the MCNOT on three copies instead of two. The three-copy protocol consists of two subprotocols. We use three identical copies of the state in each subprotocol. The output of the first subprotocol is used as input for the next. Thus, we need nine copies to run a single round. Let the three identical copies be $\rho^{(0)}$, $\rho^{(1)}$, and $\rho^{(2)}$. Subprotocol 1 ($P1$):

i. Partition the graph into two colors A and B ($V = V_A \cup V_B$ and $V_A \cap V_B = \varnothing$).

ii. Perform the MCNOT between copies $\rho^{(0)}$ and $\rho^{(1)}$ and $\rho^{(0)}$ and $\rho^{(2)}$ such that information about qubits of color A flows from $\rho^{(0)} \to \rho^{(1)}$ and $\rho^{(0)} \to \rho^{(2)}$. As a side effect information about B will flow from $\rho^{(1)}, \rho^{(2)} \to \rho^{(0)}$. See Fig. 3(a), below.

iii. Measure qubits of color A in the X basis and qubits of color B in the Z basis in states $\rho^{(1)}$ and $\rho^{(2)}$. This is a measurement of K_j for $j \in A$. If the measurement of K_j gives $+1$ (-1) we get a syndrome of 0 (1). Thus, for each $j \in A$ we have two bits of syndrome $\sigma_j^{(1)}$ and $\sigma_j^{(2)}$.

iv. Apply the correction $\prod_{j \in A} Z_j^{\sigma_j^{(1)} \cdot \sigma_j^{(2)}}$ to $\rho^{(0)}$.

For subprotocol $P2$ the roles colors A and B are interchanged.

First, we will analyze this protocol with ideal CNOT gates. This will allow us to derive simple closed-form recursion relations characterizing the behavior of the protocol, as well as analytical estimates of the threshold and efficiency. In Sec. III B we generalize to noisy gates. The analysis is restricted to density matrices that are diagonal in the graph basis (i.e., probabilistic mixtures of graph states). At the end of Sec. III B, we will show that our results are valid for arbitrary density matrices.

A. Ideal gates

Equation 3 implies that the effect of the MCNOT on $\rho^{(0)}$, $\rho^{(1)}$, and $\rho^{(2)}$ is

$$|\boldsymbol{\mu}_A^{(0)}, \boldsymbol{\mu}_B^{(0)}\rangle \mapsto |\boldsymbol{\mu}_A^{(0)}, \boldsymbol{\mu}_B^{(0)} + \boldsymbol{\mu}_B^{(1)} + \boldsymbol{\mu}_B^{(2)}\rangle \tag{4}$$

$$|\boldsymbol{\mu}_A^{(1)}, \boldsymbol{\mu}_B^{(1)}\rangle \mapsto |\boldsymbol{\mu}_A^{(1)} + \boldsymbol{\mu}_A^{(0)}, \boldsymbol{\mu}_B^{(1)}\rangle$$

$$|\boldsymbol{\mu}_A^{(2)}, \boldsymbol{\mu}_B^{(2)}\rangle \mapsto |\boldsymbol{\mu}_A^{(2)} + \boldsymbol{\mu}_A^{(0)}, \boldsymbol{\mu}_B^{(2)}\rangle\,. \tag{5}$$

Equation 2 implies that the effect of the correction is

$$|\boldsymbol{\mu}_A^{(0)}, \boldsymbol{\mu}_B^{(0)}\rangle \mapsto |\boldsymbol{\mu}_A^{(0)} + \boldsymbol{\sigma}, \boldsymbol{\mu}_B^{(0)}\rangle\,, \tag{6}$$

where $\sigma_j := \sigma_j^{(1)} \cdot \sigma_j^{(2)}$. By measuring $\rho^{(1)}$ and $\rho^{(2)}$, we get two bits of syndrome for each qubit of color A in $\rho^{(0)}$. The syndrome is conclusive; it allows us to identify, to lowest order in the error probability on which state the error occurred. We can thus do error correction instead of post-selection. This will make the protocol scale efficiently in the size of the states. The price is a reduction of the threshold value.

We now derive a recursion relation for the expectation values $\langle K_j\rangle$, $j \in 1, \ldots, N$. They yield a necessary and sufficient condition for purification. For the moment we assume that the initial state ρ is diagonal in the graph basis–i.e., ρ is a probabilistic mixture. It is then safe to consider error probabilities. This assumption is not necessary, however. It is removed in Sec. III B. Define $P_j(\rho)$ as the probability to find the eigenvalue -1 in the measurement of K_j on ρ as

$$P_j(\rho) := \mathrm{Tr}\left[\frac{1 - K_j}{2}\rho\right] = \frac{1 - \langle K_j\rangle}{2}. \tag{7}$$

Consider subprotocol $P1$. In order to analyze this protocol we make use of the fact that the error correction operation is local. It only uses information about $\langle K_j\rangle$ in each copy to apply a correction to the jth qubit in $\rho^{(0)}$. Thus, $\langle K_j\rangle$ should have nice decoupled recursion relations. We will later derive the recursion relations for the expectation value of arbitrary stabilizer elements, which in general are more complex.

First consider qubits of color B. From Eq. (4), $\mu_j^{(0)} \mapsto \mu_j^{(0)} + \mu_j^{(1)} + \mu_j^{(2)}$. Since our copies are identical, we have $P_j(\rho^{(0)}) = P_j(\rho^{(1)}) = P_j(\rho^{(2)}) = P_j$. Then, $P_j \mapsto P_j^3 + 3P_j(1 - P_j)^2$. In terms of expectation values,

$$\langle K_j\rangle' = \langle K_j\rangle^3\,. \tag{8}$$

Under concatenation of $P1$ with itself, qubits of color B are *polluted* with $\langle K_j\rangle_{\rho^{(0)}} \to \langle K_j\rangle_I = 0$.

Turning our attention to qubits of color A we note that error correction fails if $\mu_j = 1$ for more than one copy. Thus, $P_j \mapsto P_j^3 + 3P_j^2(1 - P_j)$. In terms of expectation values

$$\langle K_j\rangle' = \frac{1}{2}\left(3 - \langle K_j\rangle^2\right)\langle K_j\rangle\,. \tag{9}$$

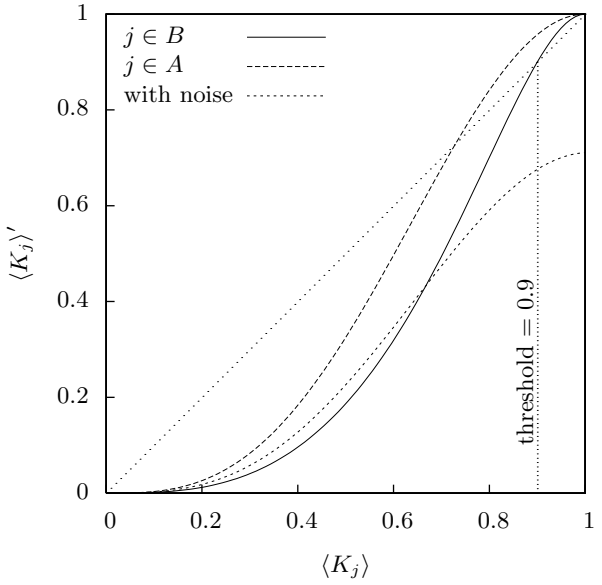

FIG. 2: Recurrence curves for the three-copy protocol. These simple curves fully encapsulate the behavior of the protocol with ideal gates. The point of intersection with $\langle K_j\rangle' = \langle K_j\rangle$ gives the threshold. If the gates are too noisy, the protocol breaks down, as indicated by the lowest curve.

Under concatenation of $P1$ with itself, qubits of color A are *purified* with $\langle K_j\rangle_{\rho^{(0)}} \to \langle K_j\rangle_{|\mathbf{0}\rangle\langle\mathbf{0}|} = 1$.

Subprotocol $P2$ is identical to $P1$ except that the roles of A and B are interchanged and the three copies are the output states from running $P1$ 3 times. The three-copy protocol is the composition of $P2$ with $P1$. Let $P = P2 \circ P1$; then Eqs. (8) and (9) imply that under the action of P

$$\langle K_j\rangle' = \begin{cases} \frac{1}{8}\left(3 - \langle K_j\rangle^2\right)^3 \langle K_j\rangle^3 & \text{if } j \in A, \\ \frac{1}{2}\left(3 - \langle K_j\rangle^6\right)\langle K_j\rangle^3 & \text{if } j \in B. \end{cases} \tag{10}$$

The recursion relations (10) have, for each color, a unique repulsive fixed point in the interval $(0, 1)$ which separates the basins of attraction for the trivial fixed point at 0 and the nontrivial fixed point at 1 (See Fig. 2). The upper fixed point corresponds to the perfect graph state. Thus, the stated protocol purifies a graph state if and only if

$$\begin{aligned} &\langle K_j\rangle > 0.7297 \text{ for all } j \text{ in } A \\ &\langle K_j\rangle > 0.9003 \text{ for all } j \text{ in } B. \end{aligned} \tag{11}$$

We can compare these thresholds to the thresholds for the post-selection protocol of [8]. For this protocol, it is not known how to derive a threshold for general noise or even probabilistic Pauli noise. However, for the particular case where only independent local phase flip errors are assumed for the initial states, recursion relations can be derived even for post-selection. Then, the P1 (post-selection) recursion relation for $\langle K_j\rangle$ with $j \in B$ is

$\langle K_j\rangle' = \langle K_j\rangle^2$ and for $j \in A$ is $\langle K_j\rangle' = \frac{2\langle K_j\rangle}{1+\langle K_j\rangle^2}$. The resulting threshold values are $\langle K_j\rangle_{\rm th} = 0.2956$ for $j \in A$ and $\langle K_j\rangle_{\rm th} = 0.5437$ for $j \in B$.

Returning to our protocol, it is possible to derive recursion relations for the expectation values of arbitrary stabilizer elements. They are not in general decoupled, but there is still a notion of locality. The generalized relation allows us to compute the recursion relations for stabilizers with small support efficiently. Define

$$K_{\boldsymbol{a},\boldsymbol{b}} := \prod_{i=1}^{|V_A|} K_i^{a_i} \prod_{j=1}^{|V_B|} K_j^{b_j}, \tag{12}$$

where $\boldsymbol{a} \in \{0,1\}^{|V_A|}$ and $\boldsymbol{b} \in \{0,1\}^{|V_B|}$. The factors in the first product are the stabilizer generators for qubits of color A, while those in the second product are for qubits of color B. Then (see Appendix A) under the action of subprotocol $P1$,

$$\langle K_{\boldsymbol{a},\boldsymbol{b}}\rangle' = \frac{1}{2^{|\boldsymbol{a}|}} \sum_{\boldsymbol{a}_1,\boldsymbol{a}_2 \ll \boldsymbol{a}} (-1)^{\boldsymbol{a}_1\cdot\boldsymbol{a}_2} \langle K_{\boldsymbol{a}+\boldsymbol{a}_1+\boldsymbol{a}_2,\boldsymbol{b}}\rangle \langle K_{\boldsymbol{a}_1,\boldsymbol{b}}\rangle \langle K_{\boldsymbol{a}_2,\boldsymbol{b}}\rangle, \tag{13}$$

where $\boldsymbol{f} \ll \boldsymbol{g}$ iff $f_j = 0$ whenever $g_j = 0$. Equations (9) and (8) are special cases for $\langle K_{\boldsymbol{a},\boldsymbol{b}}\rangle = \langle K_j\rangle$ with $j \in A, B$ respectively. An interesting feature of this equation is that it relates a correlator of weight $w = |\boldsymbol{a}| + |\boldsymbol{b}|$ to correlators of weight no more than w. This makes it feasible to calculate the recursion relations for correlators of small weight.

In order to discuss the behavior of this protocol under concatenation with itself, it is useful to switch back to probability variables. Then Eq. (10) implies that if the protocol is concatenated with itself k times,

$$P_j(\rho(k)) \le \left(\frac{P_j(\rho(0))}{P_{\rm th}}\right)^{2^k} \tag{14}$$

where $P_{\rm th}$ is the threshold error probability. The k-concatenated protocol requires 3^{2k} identical copies, thus, the protocol is exponentially efficient under concatenation. The reduction of error, Eq. (14), is not conditioned on a particular post-selected syndrome. The overhead in number of required initial states is independent of the size N of the graph state. We conclude that under concatenation the protocol reaches the reference state $|\boldsymbol{0}\rangle$ with efficient use of resources. Contrarily, for the post-selection protocol [8] the overhead acquires a dependence $\exp(\alpha N)$, with some $\alpha > 0$, due to post-selection of a particular syndrome.

B. Noisy gates

Now we investigate what happens to this protocol when the CNOT gates themselves are noisy. In the three-copy protocol CNOT gates act on the same qubit in two states $\rho^{(m)}$ and $\rho^{(n)}$. We model a noisy two-qubit gate as an ideal gate followed by the two-qubit depolarizing channel [i.e., the $SU(4)$-invariant channel]

$$\mathrm{T}^{(k)} := (1-p_2)[I] + \frac{p_2}{16}\sum_{i,j=1}^{4}\left[D_i^{(k,m)} \otimes D_j^{(k,n)}\right], \tag{15}$$

where $D_{i,j} \in \{I, X, Y, Z\}$ and k is the qubit index. $D^{(k,m)}$ acts on the kth qubit of $\rho^{(m)}$. The Z gates applied in the error-correction steps and the measurement of the syndrome are assumed to be noiseless. This is natural since the Pauli phase flips Z may be omitted as physical operations and instead accounted for in the classical syndrome processing. We will include the effect of measurement errors in the analysis when we consider the more sophisticated protocols, which have higher thresholds than the three-copy protocol. If we consider the effect of $\mathrm{T}^{(k)}$ only on $\langle K_j\rangle$ in state $\rho^{(0)}$, then using Eq. (2) we can reduce the noise to an effective error. For every $k \in V : k \in \mathcal{N}(j) \cup \{j\}$

$$\mathrm{T}_{\rm eff}^{(k,j)}(\rho^{(0)}) = \left(1 - \frac{p_2}{2}\right)[I] + \frac{p_2}{2}[Z_j^{(0)}]. \tag{16}$$

If $k \notin \mathcal{N}(j) \cup \{j\}$, then $\mathrm{T}_{\rm eff}^{(k,j)}$ is just the identity map. Since every error channel commutes with every CNOT, we can model the noisy MCNOT as the ideal MCNOT followed by $|V|$ noise channels.

The error channel Eq. (15) is local (i.e., it acts only on qubit k in $\rho^{(m)}$ and $\rho^{(n)}$). Also the error operators are Pauli operators, which map graph states to graph states, keeping ρ diagonal in the graph basis. Thus we can expect the noisy recursion relations to have the same form as Eq. (10). Considering only subprotocol $P1$, the jth qubit in $\rho^{(0)}$ is affected by $2(d+1)$ error channels. For simplicity, we assume all vertices of the graph have the same degree d. If this is not the case, then there would be a different set of recursion relations for each degree. We can then choose d to be the maximum degree, in which

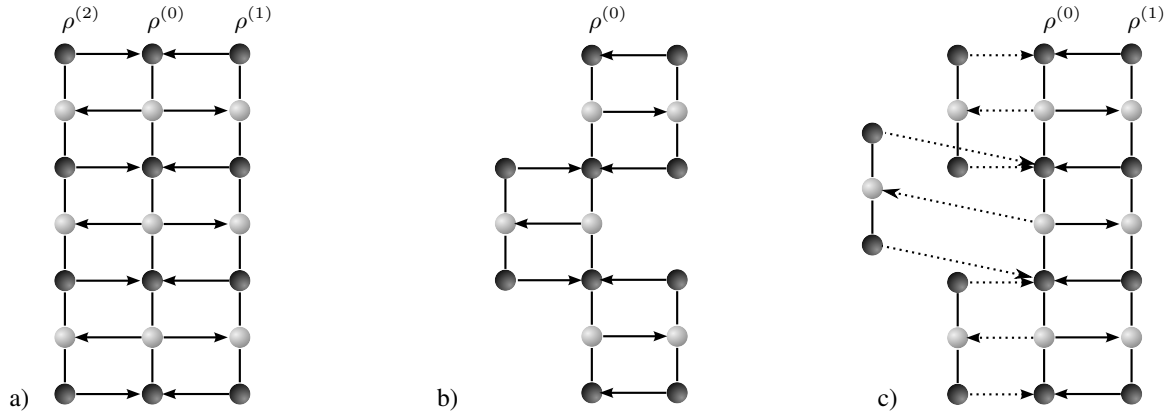

FIG. 3: The MCNOT for subprotocol P_1 in (a) The 3-copy protocol, (b) The band-aid protocol and (c) The conditional bandaid protocol. The dotted lines in c) indicate that the bandaids are applied only if there is an ambiguous syndrome at that location. Here we show graphs of degree 2, but these protocols can be applied to graphs of any degree.

case the recursion relations will be lower bounds for all other degrees. The total probability that the jth qubit is flipped by an error is $\frac{1-(1-p_2)^{2(d+1)}}{2}$. Thus, for qubits of color B,

$$\langle K_j\rangle' = \alpha^2 \langle K_j\rangle^3 , \tag{17}$$

where $\alpha = (1-p_2)^{(d+1)}$.

The situation is a little more complex for qubits of color A as the error in the MCNOT between $\rho^{(0)}$ and $\rho^{(1)}$ is propagated by the MCNOT between $\rho^{(0)}$ and $\rho^{(2)}$ (see Fig. 3(a)). However, the form of the recursion relation remains the same. We get

$$\langle K_j\rangle' = \frac{\alpha^2}{2}\left(2+\alpha^{-1}-\langle K_j\rangle^2\right)\langle K_j\rangle . \tag{18}$$

For a derivation see Appendix A 2. Composing subprotocols $P1$ and $P2$ we get the recursion relations for the three-copy protocol with noisy gates

$$\langle K_j\rangle' = \begin{cases} \frac{\alpha^8}{8}\left(2+\alpha^{-1}-\langle K_j\rangle^2\right)^3\langle K_j\rangle^3 & \text{if } j\in A, \\ \frac{\alpha^4}{2}\left(2+\alpha^{-1}-\alpha^4\langle K_j\rangle^6\right)\langle K_j\rangle^3 & \text{if } j\in B, \end{cases} \tag{19}$$

Here, qubits of color A behave worse. Solving the recursion relations for fixed points, we find that there are two non-trivial positive fixed points (see Appendix B) for $\alpha > 0.9902$. Consider the interval $[0,1]$. It has at most three fixed points $0 = f_0 < f_1 \le f_2 \le 1$. f_0 and f_2 are attractive while f_1 is repulsive. Thus f_2 will be a stable fixed point for $\alpha > 0.9902$ and $\langle K_j\rangle_{\text{initial}} > f_1$. This gives a threshold for the noise affecting the gates that scales inversely proportional to the graph degree d,

$$p_{\text{th}} \approx \frac{9.8\times 10^{-3}}{d+1}. \tag{20}$$

Specifically for degrees 2 and 4 we obtain

$$p_{\text{th}} = \begin{cases} 0.328\ \% \text{ for } d=2, \\ 0.197\ \% \text{ for } d=4. \end{cases} \tag{21}$$

This is a rather low value, but it will be substantially improved when we consider more sophisticated protocols.

We now show that the recursion relations Eq. (19) are valid regardless of whether or not the considered states are diagonal in the graph basis. To see this, let us define a depolarization operator $\mathcal{D}$ which converts an arbitrary n-qubit mixed state ρ into an n-qubit mixed state $\rho_D = \mathcal{D}\rho$ that is diagonal in the graph basis. $\mathcal{D}$ takes the form

$$\mathcal{D} = \left(\prod_{\boldsymbol{a}} \frac{[I]+[K_{\boldsymbol{a},\boldsymbol{0}}]}{2}\right)\left(\prod_{\boldsymbol{b}} \frac{[I]+[K_{\boldsymbol{0},\boldsymbol{b}}]}{2}\right), \tag{22}$$

where $\boldsymbol{a}$ and $\boldsymbol{b}$ are vectors in a basis of $\{0,1\}^{|V_A|}$ and $\{0,1\}^{|V_A|}$ respectively.

We only consider $P1$, the first round of the protocol. It is associated with a transformation $P1:\ \rho \longrightarrow \rho' = R\left(\rho^{\otimes 3}\right)$. R and $\mathcal{D}$ commute–i.e.,

$$R\left((\mathcal{D}\rho)^{\otimes 3}\right) = \mathcal{D}\circ R\left(\rho^{\otimes 3}\right), \tag{23}$$

for any ρ. For a proof see Appendix C.

Consider a recursion relation of the form

$$\langle K_{\boldsymbol{a},\boldsymbol{b}}(\rho'_D)\rangle = f_{\boldsymbol{a},\boldsymbol{b}}\left(\{\langle K_{\boldsymbol{i},\boldsymbol{j}}(\rho_D)\rangle\}\right), \tag{24}$$

with $f_{\boldsymbol{a},\boldsymbol{b}}$ some function depending on $\boldsymbol{a},\boldsymbol{b}$ as in Eq. (13). Now,

$$\begin{aligned} \langle K_{\boldsymbol{a},\boldsymbol{b}}(\rho'_D)\rangle &= \text{Tr}\left[K_{\boldsymbol{a},\boldsymbol{b}} R\left((\mathcal{D}\rho)^{\otimes 3}\right)\right] \\ &= \text{Tr}\left[K_{\boldsymbol{a},\boldsymbol{b}}\mathcal{D}\circ R\left(\rho^{\otimes 3}\right)\right] && \text{[by Eq. (23)]} \\ &= \text{Tr}\left[\mathcal{D}^\dagger(K_{\boldsymbol{a},\boldsymbol{b}})\rho'\right] && \text{(trace cyclicity)} \\ &= \langle K_{\boldsymbol{a},\boldsymbol{b}}(\rho')\rangle. && (\mathcal{D}^\dagger \equiv \mathcal{D}) \end{aligned}$$

Similarly, $\langle K_{\boldsymbol{i},\boldsymbol{j}}(\rho_D)\rangle = \langle K_{\boldsymbol{i},\boldsymbol{j}}(\rho)\rangle$, such that

$$\langle K_{\boldsymbol{a},\boldsymbol{b}}(\rho')\rangle = f_{\boldsymbol{a},\boldsymbol{b}}\left(\{\langle K_{\boldsymbol{i},\boldsymbol{j}}(\rho)\rangle\}\right). \tag{25}$$

Thus, a recursion relation of the form of Eq. (24) such as Eq. (19) holds for all states ρ and not just for diagonal states $\rho_D = \mathcal{D}\rho$.

IV. IMPROVED PROTOCOLS

A. Error model

In the following, we consider a scenario where graph states are created locally from product states, then distributed to several parties and subsequently purified. Errors occur in each of these steps–specifically, the following

- There is a two-qubit error T, Eq. (15), associated with each controlled-PHASE gate in the creation of the graph state, with probability p_2.
- A local depolarizing error with probability p_1 occurs on each graph state qubit during transmission.
- Every CNOT gate used in purification carries a two-qubit error, Eq. (15), with error probability p_2. Every measurement is modeled by a one-qubit depolarizing channel with error probability p_2 followed by a perfect measurement.

B. Bandaid protocol

In order to raise the threshold of the three-copy protocol, we will try to combine the strategies of error correction and post-selection (which has a higher threshold). One way to do this is to use small highly purified GHZ states–i.e., bandaids, to purify the graph one vertex at a time. The usual MCNOT is performed between the bandaid and the large graph state as shown in Fig. 3(b). This copies information about the central vertex into the bandaid which is then measured to give a syndrome. Since the bandaid is highly purified (for example, by post-selection), it does not pollute the large state much. It is important to note that the error correction is still local, and we expect the recursion relations to be decoupled as in the case of the three-copy protocol.

The bandaid protocol also has two subprotocols. The first one $P1$ is the following.

i. Partition the graph into two colors A and B ($V = V_A \cup V_B$ and $V_A \cap V_B = \varnothing$).

ii. The bandaids are placed over the large state such that each central qubit of the bandaid is over a vertex of qubit A for all qubits of color A. Perform the MCNOT as shown in Fig. 3(b).

iii. Measure the central qubit of each bandaid in the X basis and the other qubits in the Z basis. For each bandaid multiply the measured eigenvalues. If the product is (-1) +1 then the syndrome bit σ_j is (1) 0.

iv. Apply the correction $\prod_{j\in A} Z_j^{\sigma_j}$ to the large state.

$P2$ is the same as $P1$, with the roles of colors A and B reversed.

Consider subprotocol $P1$. For qubits of color B the argument is very similar to the three-copy protocol, except that each qubit is affected by two gates from each of d bandaids. Thus,

$$\langle K_j\rangle' = (1-p_2)^{2d}\,\langle K_j\rangle\,\langle K_j\rangle_b^d\,, \tag{26}$$

where $\langle K_j\rangle_b$ is the constant initial purity of the bandaid.

For qubits of color A, first suppose that the CNOT gates are ideal. Then, a simple transfer of purity occurs:

$$\langle K_j\rangle' = \langle K_j\rangle_b\,. \tag{27}$$

If the gates are noisy, Eq. (27) is multiplied by a noise factor of the form $(1-p_2)^{f(d)}$ as in the case of the three-copy protocol. There is a subtlety involving the temporal ordering of the bandaids. The bandaids do not all commute with each other. There are $1+d(d-1)$ bandaids that affect qubit j. One of them is the bandaid that is used to purify the qubit. On average $k = \frac{d(d-1)}{2}$ of the rest will be applied before the purifying one. Any effect from the k prior bandaids will be erased by the purifying bandaid [see Eq. (27)]. The purifying bandaid has $d+1$ noisy CNOTs affecting $\langle K_j\rangle$; since the noisy MCNOT is modeled as an ideal MCNOT followed by noise, no information about the noise is propagated to the bandaid. Thus, the noise will commute with the error correction procedure. Since a measurement error that flips the central qubit of the bandaid will cause us to apply the wrong error correction operator, it can also be reduced to an effective error as given by Eq. (16). Thus, $f(d) = 2(d+1)+k$ and we have

$$\langle K_j\rangle' = (1-p_2)^{\frac{d(d+3)+4}{2}}\,\langle K_j\rangle_b\,. \tag{28}$$

Combining subprotocols $P1$ and $P2$, we get the recursion relations for the bandaid protocol with noisy gates as well as noisy measurements

$$\langle K_j\rangle' = \begin{cases} (1-p_2)^{\frac{d(d+7)+4}{2}}\,\langle K_j\rangle_b^{d+1} & \text{for } j\in A,\\ (1-p_2)^{\frac{d(d+3)+4}{2}}\,\langle K_j\rangle_b & \text{for } j\in B. \end{cases} \tag{29}$$

The behavior of qubits of color A is worse, and we will use their purity as the final purity of the large state.

As per our error model in Sec. IV A, the noisy CPHASE, CNOT and measurement gates are parametrized by p_2. The noisy transmission channel is parametrized by p_1. For the final result, we need to know the quality of the bandaids. We assume that these are also created locally, then transmitted and purified. The bandaids, however, are of fixed size and may thus be purified by the post-selection protocol [9] with the higher threshold. The output quality of the purified bandaids is, to leading order in p_2,

$$\langle K_j\rangle_b = 1-(d+1)p_2, \tag{30}$$

such that

$$\langle K_j\rangle = 1-\frac{1-d(3d+11)+6}{2}p_2, \tag{31}$$

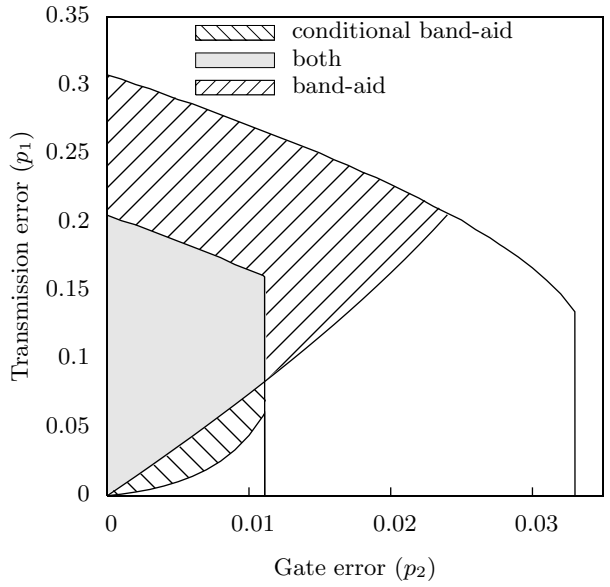

FIG. 4: Trade-off curves for the bandaid and conditional bandaid protocols ($d = 4$). The decreasing curves represent the breakdown of the post-selection protocol, when there is too much error. The increasing curves demarcate the region where the final purity of the purified states is higher than the purity of the unpurified states. It makes sense to purify in the shaded regions.

for small p_2 [from Eqs. (29) and (30)]. As Eq. (31) shows, with increasing graph degree the effect of errors in the purification process is strongly enhanced. One may therefore ask the question whether it is useful to purify at all or whether the transmitted state should be used right away. To decide this we compute $\langle K_j \rangle$ after graph-state creation and transmission,

$$\langle K_j \rangle = (1-p_2)^{\frac{d(d+1)}{2}} (1-p_1)^{d+1}. \qquad (32)$$

See Appendix D for a derivation. We compare this expression with Eq. (29) and find that there is indeed a parameter region where it makes sense to purify. This region is displayed for graphs of degree $d = 4$ in Fig. 4. It is bounded from above and right by the curve which indicates the breakdown of the bandaid purification according to the post-selection protocol [9]. If we use post-selection to obtain bandaids of high purity, then the threshold of the bandaid protocol for degree d graph states equals the threshold for purification of a $d+1$-qubit GHZ state with the post-selection protocol [8]. However, the output purity of the bandaid protocol is smaller. Only above the ascending curve is it advantageous to purify.

C. Conditional bandaid protocol

In order to correct the d^2 dependence of the fixed point in the bandaid protocol, we will combine it with the three-copy protocol. The hybrid protocol, called the conditional bandaid protocol, sacrifices in threshold to improve the fixed point. The fixed-point behavior, at least to linear order in gate noise, is almost as good as that of the post-selection protocol.

This protocol proceeds in the same fashion as the three-copy protocol, except that two copies are used per round, and wherever a measurement of K_j yields eigenvalue -1 (i.e., an error), a post-selected bandaid is applied to purify qubit j (see Fig. 3). For small gate noise, we expect to have to apply only a few bandaids per round, nonetheless, the threshold is set by the qubits to which we have to apply bandaids. Locations where a measurement of K_j yields 1 are error free to lowest order. Once again, we have two subprotocols, P_1 and P_2 each purifying a different color.

The analysis is similar to that used in arriving at Eq. (19) for the three-copy protocol. However, the situation is complicated by the fact that the bandaids are applied conditioned on the results of measuring $\rho^{(1)}$. As a result, the recursion relations for the one point correlators are no longer completely decoupled. We can, however, find a simple lower bound on them.

Define $\langle K_b \rangle$ to be the minimum purity of the post-selected bandaid. It is a constant. For simplicity we assume that all qubits in the bandaid have this purity. As before, we assume that the graph of the large state is translationally invariant, i.e., all vertices have the same degree. The definition $\beta \equiv (1-p_2)^2 \langle K_b \rangle$ will be useful. Consider qubits of color A in subprotocol P_1, then, by a derivation similar to Eq. (18),

$$\langle K_j \rangle' = \frac{\alpha}{2} \left(2\alpha \langle K_j \rangle + \langle K_b \rangle - \alpha \langle K_b \rangle \langle K_j \rangle^2 \right), \qquad (33)$$

where $\alpha = (1-p_2)^{d+1}$ as before.

So far, we have been exact. Now consider subprotocol P_2. Again focus on qubits of color A. Break P_2 down into two steps. In step one, we apply the MCNOT to $\rho^{(0)}$ and $\rho^{(1)}$. It can be readily verified that $\langle K_j \rangle \mapsto \alpha \langle K_j \rangle^2$. In step two, bandaids conditioned on the measurement outcome are applied to qubits of color B. Let $\boldsymbol{y} \in \{0,1\}^d$ be the measurement results for the neighbors of qubit j. A measurement result of one means a bandaid must be applied at that location. If a bandaid is applied to a neighbor of j, $\langle K_j \rangle$ is affected by the errors on the bandaid, characterized by $\langle K_b \rangle$ and by two noisy CNOTs. Thus $\langle K_j \rangle \mapsto \beta^{|\boldsymbol{y}|} \langle K_j \rangle$. Summing over measurement outcomes and including step one, we get

$$\langle K_j \rangle' = \alpha \left(\sum_{k=0}^{d} \left(\sum_{|\boldsymbol{y}|=k} q_{\boldsymbol{y}} \beta^k \right) \right) \langle K_j \rangle^2, \qquad (34)$$

where $q_{\boldsymbol{y}}$ is the probability of measurement outcome $\boldsymbol{y}$. Unfortunately, $q_{\boldsymbol{y}}$ is a function of the general stabilizer expectation values $\langle K_{\boldsymbol{a},\boldsymbol{b}} \rangle$, so we will resort to finding a lower bound. Since $q_0 = 1 - \sum_{\boldsymbol{y} \neq 0} q_{\boldsymbol{y}}$, we can rewrite the

above equation as

$$\langle K_j\rangle' = \alpha_a \left((1 - \sum_{\boldsymbol{y}\neq 0} q_{\boldsymbol{y}}) + \sum_{k>0}^{d} \sum_{|\boldsymbol{y}|=k} q_{\boldsymbol{y}} \beta^k \right) \langle K_j\rangle^2$$
$$\geq \alpha \left(1 - (1-\beta^d) \sum_{\boldsymbol{y}\neq 0} q_{\boldsymbol{y}} \right) \langle K_j\rangle^2 ,$$

using $\beta \leq 1$ to arrive at the inequality.

Now, q_0 is just the probability that no error is detected on any of the neighbors of j. Let p_j be the probability of detecting an error on site j. Then, by definition, $\langle K_i\rangle = 1 - 2\sum_{y|y_i=1} q_y$. This implies that $\sum_{\boldsymbol{y}\neq 0} q_{\boldsymbol{y}} \leq \sum_{i\in\mathcal{N}(j)} \frac{1-\langle K_i\rangle}{2}$. Putting this into the above inequality,

$$\langle K_j\rangle' \geq \alpha \left(1 - \frac{d}{2}(1-\beta^d)\left(1 - \langle K_i\rangle\right) \right) \langle K_j\rangle^2 , \qquad (35)$$

where $\langle K_i\rangle$ is the purity of qubits of color B from the previous round.

Solving for the fixed point, we get, to leading order in gate noise p_2,

$$\langle K_j\rangle = 1 - 2(d+1)p_2. \qquad (36)$$

Comparing this to Eq. (30), we see that the fixed-point scaling with degree is almost as good as in the post-selection protocol. We now apply the conditional bandaid protocol to the same situation–of a graph state being shared among widely separated parties, as for the bandaid protocol. The results for a degree four state are plotted in Fig. 4. We see that the threshold (upper) curve is worse, whereas the fixed-point (lower) curve is better for this protocol, as compared to the bandaid protocol. The total purifiable area is smaller, indicating that it breaks down faster. In some sense, we have traded threshold for fixed point. These conclusions hold for arbitrary degree, and the curves are independent of the size of the state, making this protocol eminently suitable for the purification of large bi-colorable graph states.

V. CONCLUSION AND OUTLOOK

We have described novel purification protocols for bi-colorable graph states and discussed their performance. The criteria for our protocols are that they do not break down in the presence of small amounts of noise in the purification process, that they have a high purification threshold and good output quality, scale efficiently, and be analytically tractable.

Our final protocol can, for relevant graph states of degree 4, tolerate 1% gate or 20% local transmission error. These are about 1/3 and 2/3 of the respective values for the post-selection protocol [8, 9]. However, in contrast to this reference protocol, our protocol scales efficiently with the graph size.

All our protocols can be treated analytically. In particular, for the three-copy protocol we derive closed, exact one-dimensional recursion relations in the appropriate observables, irrespective of the size of the state.

We would like to comment on the influence of the graph degree for the purification threshold. First note that for the three-copy protocol of Section III, in the case of perfect purification gates, the recursion relations (10) are completely independent of the graph structure, and so are the thresholds (11). This behavior changes if noise is included in the purification. The critical noise level per purification gate—at which the protocol breaks down—scales inversely proportional with the graph degree. The unfavorable dependence on the graph degree is present in all three protocols we discuss. Thus, the lesson we learn for the case of noisy purification is to beware of large graph degrees. Large graph degrees occur, for example, in graphs states corresponding to codewords of concatenated CSS codes.

We would also like to comment on the structure of the nontrivial fixed point in our protocols. In the case of erroneous purification gates, the nontrivial fixed point is not completely specified by the lowest order expectation values $\langle K_j\rangle$ and it remains to be discussed which error correlations are removed by the purification protocol. As a first result in this direction, for the three-copy protocol discussed in Section III we have shown (in Appendix A 3) that correlations of stabilizer expectation values located on non-overlapping supports are not introduced by the purification procedure if they are absent initially. This implies that such correlations are absent in all purified states which end up at the same fixed point as the perfect state. We show in Appendix B that the fixed point for two-generator correlations with distinct support is unique, which is enough to establish the result that all states at the fixed point obey the relation $\langle K_i K_j\rangle = \langle K_i\rangle \langle K_j\rangle$ for such correlations.

A question of further interest is whether the nontrivial fixed point of the protocol is unique at all levels of correlations. This would imply $\langle K_{\boldsymbol{i}+\boldsymbol{j}}\rangle = \langle K_{\boldsymbol{i}}\rangle \langle K_{\boldsymbol{j}}\rangle$ for all correlations with distinct supports.

Another question of further interest is whether the described or related protocols may be used to boost the threshold value for fault-tolerant quantum computation [20, 21, 22, 23, 24] based on graph states.

Acknowledgments

We would like to thank John Preskill, Frank Verstraete, Jiannis Pachos, Maarten van den Nest, Eric Hostens, Akimasa Miyake, Wolfgang Dür, Simon Anders, Hans Briegel, Panos Aliferis and Krittika Kanjilal for useful discussions. K.G. is supported by DOE Grant No. DE-FG03-92-ER40701. R.R. has been supported at Caltech by MURI under Grant No. DAAD19-00-1-0374 and by the National Science Foundation under Contract No. PHY-0456720, and is supported by the Government

of Canada through NSERC and by the Province of Ontario through MEDT. Additional support was provided by the National Science Foundation under Grant No. PHY99-07949 during the workshop "Topological Phases and Quantum Computation" at the KITP Santa Barbara, and by the Austrian Academy of Sciences.

APPENDIX A: GENERALIZED RECURSION RELATIONS

We now derive the generalized recursion relations [Eq. (13)] for the three-copy protocol. While the method used for this derivation is less intuitive, it yields recursion relations for arbitrary stabilizer elements and can handle noisy gates easily.

1. Noiseless gates

In order to derive Eq. (13) we work in the *stabilizer basis*. Because $\rho^{(0)}$ is diagonal and the set $\{\langle K_{\boldsymbol{a},\boldsymbol{b}}\rangle\}$ where $\boldsymbol{a} \in \{0,1\}^{|V_A|}, \boldsymbol{b} \in \{0,1\}^{|V_A|}$ forms a complete set of observables, we can write an expansion $\rho^{(0)} = \frac{1}{2^{|V_A|+|V_B|}} \sum_{\boldsymbol{a},\boldsymbol{b}} \langle K_{\boldsymbol{a},\boldsymbol{b}}\rangle K_{\boldsymbol{a},\boldsymbol{b}}$.

Consider subprotocol P_1, which purifies the A subgraph. The initial state is $\rho^{(0)} \otimes \rho^{(1)} \otimes \rho^{(2)}$, which can be rewritten as a sum over $\boldsymbol{a}, \boldsymbol{b}$ of terms of the form

$$\langle K_{\boldsymbol{a}^{(0)},\boldsymbol{b}^{(0)}}\rangle \langle K_{\boldsymbol{a}^{(1)},\boldsymbol{b}^{(1)}}\rangle \langle K_{\boldsymbol{a}^{(2)},\boldsymbol{b}^{(2)}}\rangle \times K_{\boldsymbol{a}^{(0)},\boldsymbol{b}^{(0)}} K_{\boldsymbol{a}^{(1)},\boldsymbol{b}^{(1)}} K_{\boldsymbol{a}^{(2)},\boldsymbol{b}^{(2)}}. \quad \text{(A1)}$$

The protocol is linear, so we track the evolution of each term seperately. Performing step (ii), this term becomes

$$\langle K_{\boldsymbol{a}^{(0)},\boldsymbol{b}^{(0)}}\rangle \langle K_{\boldsymbol{a}^{(1)},\boldsymbol{b}^{(1)}}\rangle \langle K_{\boldsymbol{a}^{(2)},\boldsymbol{b}^{(2)}}\rangle \times K_{\boldsymbol{a}^{(0)}+\boldsymbol{a}^{(1)}+\boldsymbol{a}^{(2)},\boldsymbol{b}^{(0)}} K_{\boldsymbol{a}^{(1)},\boldsymbol{b}^{(0)}+\boldsymbol{b}^{(1)}} K_{\boldsymbol{a}^{(2)},\boldsymbol{b}^{(0)}+\boldsymbol{b}^{(2)}}. \quad \text{(A2)}$$

Now consider step (iii). Suppose we get measurement outcomes $\boldsymbol{\lambda}^{(1)}, \boldsymbol{\lambda}^{(2)}$ for the stabilizers in subgraph A on copies $\rho^{(1)}, \rho^{(2)}$. Then the resultant state is given by applying the projector

$$\frac{1}{2^{2|V_A|}} \prod_{j=1}^{|V_A|} [I] \otimes \left([I] + (-1)^{\lambda_j^{(1)}} K_j^{(1)}\right) \otimes \left([I] + (-1)^{\lambda_j^{(2)}} K_j^{(2)}\right). \quad \text{(A3)}$$

All the single-site operators involved commute, so this term is a product of stabilizers in $\boldsymbol{b}$ and terms of the form

$$\left([I] + (-1)^{\lambda_j^{(k)}} K_j^{(k)}\right) \left(K_j^{(k)}\right)^{a_j^{(k)}}.$$

Here $k = 1, 2$. Discarding $\rho^{(1)}, \rho^{(2)}$, we perform a partial trace over these systems (recalling that $K_{\boldsymbol{a},\boldsymbol{b}}$ are all traceless except $K_{\boldsymbol{0},\boldsymbol{0}} = I$). In the above term, only the coefficient of $[I]$ contributes, which is $(-1)^{\lambda_j^{(k)} a_j^{(k)}}$. Including the stabilier operator, we are left with

$$\delta_{\boldsymbol{b}^{(0)},\boldsymbol{b}^{(1)}} \delta_{\boldsymbol{b}^{(1)},\boldsymbol{b}^{(2)}} \frac{(-1)^{\boldsymbol{\lambda}^{(1)}\cdot\boldsymbol{a}^{(1)} + \boldsymbol{\lambda}^{(2)}\cdot\boldsymbol{a}^{(2)}}}{2^{2|V_A|}} \times K_{\boldsymbol{a}^{(0)}+\boldsymbol{a}^{(1)}+\boldsymbol{a}^{(2)},\boldsymbol{b}^{(1)}}, \quad \text{(A4)}$$

where $\delta_{\boldsymbol{p},\boldsymbol{q}}$ is the Kronecker delta on each component of $\boldsymbol{p}, \boldsymbol{q}$. Note that we must have $\boldsymbol{b}^{(0)} = \boldsymbol{b}^{(1)} = \boldsymbol{b}^{(2)}$ or the term is zero.

Now examine the action of the Pauli $[Z]$ operator in this basis. $[Z]K_{\boldsymbol{a},\boldsymbol{b}} = ZK_{\boldsymbol{a},\boldsymbol{b}}Z = -1^k K_{\boldsymbol{a},\boldsymbol{b}}$, where $k = 0$ iff Z and $K_{\boldsymbol{a},\boldsymbol{b}}$ commute. Effectively, Z is a diagonal matrix with entries ± 1. Identical reasoning applies to X and Y. This will make it very easy to add gate noise into the analysis. It also allows us to say that the net effect of the error-correction step iv is to multiply Eq. (A4) by a factor of $(-1)^{(\boldsymbol{\lambda}^{(1)}\times\boldsymbol{\lambda}^{(2)})\cdot(\boldsymbol{a}^{(0)}+\boldsymbol{a}^{(1)}+\boldsymbol{a}^{(2)})}$, where $(\boldsymbol{p}\times\boldsymbol{q})_j \equiv p_j \cdot q_j$. To simplify the notation, change the basis to $\boldsymbol{a} \equiv \boldsymbol{a^{(0)}} + \boldsymbol{a^{(1)}} + \boldsymbol{a^{(2)}}$, $\boldsymbol{b} \equiv \boldsymbol{b}^{(0)}$. Then the term becomes

$$\delta_{\boldsymbol{b},\boldsymbol{b}^{(1)}} \delta_{\boldsymbol{b},\boldsymbol{b}^{(2)}} \frac{(-1)^{\boldsymbol{\lambda}^{(1)}\cdot\boldsymbol{a}^{(1)} + \boldsymbol{\lambda}^{(2)}\cdot\boldsymbol{a}^{(2)} + (\boldsymbol{\lambda}^{(1)}\times\boldsymbol{\lambda}^{(2)})\cdot\boldsymbol{a}}}{2^{2|V_A|}} K_{\boldsymbol{a},\boldsymbol{b}}.$$

In this notation and ignoring the delta functions, the original coefficient in Eq. (A2) is $\langle K_{\boldsymbol{a}+\boldsymbol{a^{(1)}}+\boldsymbol{a^{(2)}},\boldsymbol{b}}\rangle \langle K_{\boldsymbol{a^{(1)}},\boldsymbol{b}}\rangle \langle K_{\boldsymbol{a^{(2)}},\boldsymbol{b}}\rangle$. We will now get conditions under which this term contributes to the coefficient of $K_{\boldsymbol{a},\boldsymbol{b}}$.

Summing over measurement outcomes, the coefficient of $K_{\boldsymbol{a},\boldsymbol{b}}$ is

$$\langle K_{\boldsymbol{a}+\boldsymbol{a^{(1)}}+\boldsymbol{a^{(2)}},\boldsymbol{b}}\rangle \langle K_{\boldsymbol{a^{(1)}},\boldsymbol{b}}\rangle \langle K_{\boldsymbol{a^{(2)}},\boldsymbol{b}}\rangle \times \frac{1}{2^{2|V_A|}} \sum_{\boldsymbol{\lambda}^{(1)},\boldsymbol{\lambda}^{(2)}} (-1)^{\boldsymbol{\lambda}^{(1)}\cdot\boldsymbol{a}^{(1)} + \boldsymbol{\lambda}^{(2)}\cdot\boldsymbol{a}^{(2)} + (\boldsymbol{\lambda}^{(1)}\times\boldsymbol{\lambda}^{(2)})\cdot\boldsymbol{a}}.$$

The sum can be reexpressed as

$$\prod_{j=1}^{|V_A|} \sum_{\lambda_j^{(1)},\lambda_j^{(2)}=0}^{1} (-1)^{\lambda_j^{(1)} a_j^{(1)} + \lambda_j^{(2)} a_j^{(2)} + \lambda_j^{(1)} \lambda_j^{(2)} a_j}.$$

If $a_j = 0$, then the jth factor is zero unless $a_j^{(1)} = a_j^{(2)} = 0$, in which case it is 4. Hence, for the term $\langle K_{\boldsymbol{a}+\boldsymbol{a^{(1)}}+\boldsymbol{a^{(2)}},\boldsymbol{b}}\rangle \langle K_{\boldsymbol{a^{(1)}},\boldsymbol{b}}\rangle \langle K_{\boldsymbol{a^{(2)}},\boldsymbol{b}}\rangle$ to survive the procedure, we must have $\boldsymbol{a}^{(1)}, \boldsymbol{a}^{(2)} \ll \boldsymbol{a}$. If this holds, then an overall factor of $4^{|V_A|-|\boldsymbol{a}|}$ comes out. If $a_j = 1$, then a straightforward calculation shows that the jth factor contributes a factor of $2(-1)^{a_j^{(1)} a_j^{(2)}}$. The overall numerical factor is thus $\frac{1}{2^{|\boldsymbol{a}|}}$. To get the new value of $\langle K_{\boldsymbol{a},\boldsymbol{b}}\rangle$, we simply sum over $\boldsymbol{a}^{(1)}, \boldsymbol{a}^{(2)}$ since these and only these will contribute to the support of $K_{\boldsymbol{a},\boldsymbol{b}}$ under P1. This gives Eq. (13).

2. Noisy gates

Adding noise to the gates requires very little additional work. We can rewrite the depolarizing channel on qubit j of copy k as

$$\begin{aligned} D_j^{(k)}[\rho] &= \frac{1}{2}\left([I] + [Z]_j^{(k)}\right)\frac{1}{2}\left([I] + [X]_j^{(k)}\right)[\rho] \\ &\equiv P_Z^{(j,k)} P_X^{(j,k)}. \end{aligned}$$

It was shown above that $[X], [Z]$ have ± 1 on the diagonal. Thus writing the noise channel in this form illustrates how the noise components act as projectors $P_{Z_j}^{(k)}$, $P_{X_j}^{(k)}$. If a specific ket is affected by noise on site j of copy k, it will be an eigenvector of $D_j^{(k)}$ with zero eigenvalue.

The noise from a CNOT at site j between copies i and k is

$$E_j^{(i),(k)} \equiv (1-p_2) + p_2 (P_{Z_j}^{(i)} P_{X_j}^{(i)})(P_{Z_j}^{(k)} P_{X_j}^{(k)}). \tag{A5}$$

If a ket $K_{\boldsymbol{a},\boldsymbol{b}}$ is affected by any of these noise terms (that is, if the noise anticommutes with $K_{\boldsymbol{a},\boldsymbol{b}}$), it will be projected to zero and thus acquire a $(1-p_2)$ multiplier overall.

The noise from the first MCNOT is $E_{01} \equiv \prod_j E_j^{(0),(1)}$ and from the second MCNOT is $E_{02} \equiv \prod_j E_j^{(0),(2)}$. Clearly the overall multiplier is independent of the measurement outcomes, so the analysis for Eq. (A3) still holds. The recursion relations are then similar in structure to Eq. (13), except that coefficients dependent on $(1-p_2)$ are inserted before each term.

We illustrate this by calculating the recursion relations for $\langle K_j \rangle$. If $j \in B$, there is no sum in Eq. (13), and $\langle K_j \rangle \to E_j \langle K_j \rangle^3$. The only noise terms that anticommute with K_j [and hence give factors of $(1-p_2)$] are those in $j \cup N_j$. There are $2(d+1)$ of these (since there are two sets of noisy gates), so $\langle K_j \rangle \to (1-p_2)^{2(d+1)} \langle K_j \rangle^3$, which is Eq. (17).

Now suppose $j \in A$. Let $\boldsymbol{j} = (0, \ldots, 0, j, 0, \ldots, 0)$. Our sum is over $\boldsymbol{a}^{(1)}, \boldsymbol{a}^{(2)} \in \{\boldsymbol{0}, \boldsymbol{j}\}$, and $\boldsymbol{b} = \boldsymbol{0}$. Since we are interested only in $\langle K_j \rangle$, our effective noise model is $[X]_k \mapsto [Z]_j \forall k \in N_j$ and $[X]_j \mapsto [I]$. All other noise terms do not affect the state. Then

$$E_{01} \mapsto \left[(1-p_2) + p_2 P_Z^{(j,0)} P_Z^{(j,1)}\right]^{d+1}. \tag{A6}$$

A similar replacement holds for E_{02}. E_{01} acts on terms $K_{\boldsymbol{j}+\boldsymbol{a}^{(1)}+\boldsymbol{a}^{(2)},\boldsymbol{0}} K_{\boldsymbol{a}^{(1)},\boldsymbol{0}}$ and gives a factor of 1 iff $\boldsymbol{j} + \boldsymbol{a}^{(1)} + \boldsymbol{a}^{(2)} = \boldsymbol{0}, \boldsymbol{a}^{(1)} = \boldsymbol{0} \Rightarrow \boldsymbol{j} = \boldsymbol{a}^{(2)}, \boldsymbol{a}^{(1)} = \boldsymbol{0}$, and a factor of $(1-p_2)^{d+1}$ otherwise.

Performing the MCNOT between $\rho^{(0)}$ and $\rho^{(2)}$, the noise channel E_{02} acts on the kets $K_{\boldsymbol{j}+\boldsymbol{a}^{(1)},\boldsymbol{0}} K_{\boldsymbol{a}^{(2)}}$, which gives a factor of 1 iff $\boldsymbol{j} + \boldsymbol{a}^{(1)} = \boldsymbol{0}, \boldsymbol{a}^{(2)} = \boldsymbol{0}$ and $(1-p)^{d+1}$ otherwise. Putting in each of the four cases $a_j^{(1)}, a_j^{(2)} \in \{0, 1\}$ gives us Eq. (18).

3. Behavior of correlations

If we take two qubits j, k such that $\mathcal{N}(j) \cap \mathcal{N}(k) = \oslash$, then the noise terms on sites in $\mathcal{N}(k) \cup k$ do not affect terms involving j and vice versa. Hence the sum over terms in the recursion relation for $\langle K_{jk} \rangle$ will factor into $\langle K_j \rangle \langle K_k \rangle$. If initially $\langle K_j K_k \rangle = \langle K_j \rangle \langle K_k \rangle$, then the three-copy protocol will not generate any new correlations between these regions.

APPENDIX B: UNIQUENESS OF THE FIXED POINT

Here we show that the three-copy protocol has a unique fixed point for stabilizer elements $\langle K_{\boldsymbol{a},\boldsymbol{b}} \rangle$ with weight $w = |\boldsymbol{a}| + |\boldsymbol{b}| \leq 2$. The recursion relations for stabilizer elements of weight $w > 1$ [see Eq. (13)] depend only on stabilizer elements whose weight is at most w. Thus, we can use an inductive argument. If all the stabilizer elements of weight less than w have reached a fixed point, they become constants and then the recursion relation for elements of weight w will have the same form as those for weight one (i.e., they will depend only on stabilizer elements of weight w). First consider the case when $|\boldsymbol{a}|, |\boldsymbol{b}| \leq 1$. For this case, the three-copy recursion relations Eq. (13) have the form

$$\begin{aligned} f(z) &= az + bz^3, \\ g(z) &= cz + dz^3, \end{aligned}$$

with $a, c > 0$ and $bd < 0$. The presence of noise does not change the form of the recursion relations, it only multiplies each term by a number between 0 and 1 (see Appendix A 2). Let $y = z^2$ and $x = dy + c$. Define

$$p(x) := f(g(z))/z - 1 = bx^4 - bcx^3 + adx - d.$$

The signature of $p(x)$ is

$$\begin{aligned} p(x) &: - + + -, \\ p(-x) &: - - - - . \end{aligned}$$

Then by Descartes' rule of signs [25], $p(x)$ has at least two complex roots. Thus the recursion relation $f(g(z)) = z$ has at most two positive fixed points. The recursion relation $g(f(z)) = z$ can be analyzed identically. It was already argued in Sec. III B that this means that there is a unique attractive fixed point.

Now consider the case $|\boldsymbol{a}| = 2$ and $|\boldsymbol{b}| = 0$. The recursion relations now have the form

$$\begin{aligned} f(z) &= az^3 + bz + c \\ g(z) &= dz^3. \end{aligned}$$

It is easily checked that a, c, and d are positive. The sign of b is harder to fix, but we note that for there to be a fixed point at all, b must be negative. The

case $f(g(z)) = z$ is easily analyzed, as above, to show that there are at most two positive roots. Let $p(z) = g(f(z))$. To conclude the proof we need two technical results. (i) If the smallest support expectation value $\langle K_a \rangle$ has reached its fixed point value $\langle K_a \rangle_{\rm fp}$ then the physically allowed values for $\langle K_{a+a'} \rangle$ form the interval $I = [2\langle K_a \rangle_{\rm fp} - 1, 1]$. (ii) $f(z) \geq 0$ for all $z \in I$. Proof of (i) (a) z allowed $\Rightarrow z \in I$: $P = \frac{1-K_a}{2}\frac{1-K_{a'}}{2}$, with $a \neq a'$, is a projector, hence $\langle P \rangle \geq 0$. Thus $z = \langle K_{a+a'} \rangle \geq \langle K_a \rangle + \langle K_{a'} \rangle - 1$ (*). Evaluate (*) at fixed point $\langle K_a \rangle_{\rm fp}$. $z \leq 1$ is obvious. (b) $z \in I \Rightarrow z$ allowed: For an initial state of the protocol, interpolate between $\rho_1 = \langle K_a \rangle_{\rm fp} \rho_{++} + (1 - \langle K_a \rangle_{\rm fp})/2\,(\rho_{+-} + \rho_{-+})$ and $\rho_2 = \langle K_a \rangle_{\rm fp} \rho_{++} + (1 - \langle K_a \rangle_{\rm fp})\rho_{--}$. (The signs "$\pm$" refer to the eigenvalues of K_a and $K_{a'}$, respectively.) Proof of (ii). Be $\langle K_a \rangle_{\rm fp}, \langle K_b \rangle_{\rm fp} > 0$ and $z \in I$. Assume as an hypothesis $f(z) < 0$. Apply (*) to the state after application of P1, at the fixed point $\langle K_a \rangle_{\rm fp}$, $\forall a \in A$. Hence $0 \geq f(z) \geq 2\langle K_b \rangle_{\rm fp} - 1$. (Under P1 the fixed point value $\langle K_a \rangle_{\rm fp}$ for $a \in A$ is mapped to $\langle K_b \rangle_{\rm fp}$ for $b \in B$, assuming all vertices have the same degree.) Thus, $\langle K_b \rangle_{\rm fp} \leq 1/2$. But then $\langle K_b \rangle_{\rm fp} = 0$, which is a contradiction. Hence $f(z) \geq 0$.

Now, $p''(z) = g''(f(z))f'(z)^2 + g'(f(z))f''(z)$ such that, with (ii), $p'' \geq 0$ for all $z \in I$. Thus, $p(z)$ is convex on I. With (i), I is a single interval such that $p(z)$ and z intersect at most twice in I. At most one of these fixed points is attractive.

APPENDIX C: THE DEPOLARIZING OPERATOR

In order to prove that the depolarizing operator $\mathcal{D}$ defined in Eq. (22) commutes with the evolution operator $R = {\rm Tr}_{(1,2)}\, M \circ \mathcal{E} \circ U$, we note that the protocol step P1 consists of a unitary part U, an error channel $\mathcal{E}$ comprising probabilistic Pauli errors, and a measurement ${\rm Tr}_{(1,2)}\, M$, where M is a projector. U consists of a set of transversal CNOT-gates and acts on the stabilizer as

$$\begin{aligned} K^{(0)}_{\boldsymbol{a},\boldsymbol{b}} &\longrightarrow K^{(0)}_{\boldsymbol{a},\boldsymbol{b}} K^{(1)}_{\boldsymbol{a},\boldsymbol{0}} K^{(2)}_{\boldsymbol{a},\boldsymbol{0}}, \\ K^{(1)}_{\boldsymbol{a},\boldsymbol{b}} &\longrightarrow K^{(1)}_{\boldsymbol{a},\boldsymbol{b}} K^{(0)}_{\boldsymbol{0},\boldsymbol{b}}, \\ K^{(2)}_{\boldsymbol{a},\boldsymbol{b}} &\longrightarrow K^{(2)}_{\boldsymbol{a},\boldsymbol{b}} K^{(0)}_{\boldsymbol{0},\boldsymbol{b}}. \end{aligned} \tag{C1}$$

Now note that $\frac{[I]+[K^{(1)}_{\boldsymbol{0},\boldsymbol{b}} K^{(0)}_{\boldsymbol{0},\boldsymbol{b}}]}{2}\frac{[I]+[K^{(0)}_{\boldsymbol{0},\boldsymbol{b}}]}{2}\frac{[I]+[K^{(2)}_{\boldsymbol{0},\boldsymbol{b}} K^{(0)}_{\boldsymbol{0},\boldsymbol{b}}]}{2} = \frac{[I]+[K^{(0)}_{\boldsymbol{0},\boldsymbol{b}}]}{2}\frac{[I]+[K^{(1)}_{\boldsymbol{0},\boldsymbol{b}}]}{2}\frac{[I]+[K^{(2)}_{\boldsymbol{0},\boldsymbol{b}}]}{2}$ etc., such that

$$U \circ \mathcal{D}^{(0)}\mathcal{D}^{(1)}\mathcal{D}^{(2)} = \mathcal{D}^{(0)}\mathcal{D}^{(1)}\mathcal{D}^{(2)} \circ U. \tag{C2}$$

The operations $\mathcal{D}^{(0)}\mathcal{D}^{(1)}\mathcal{D}^{(2)}$ and $\mathcal{E}$ commute because both are linear combinations of Pauli superoperators,

$$\mathcal{E} \circ \mathcal{D}^{(0)}\mathcal{D}^{(1)}\mathcal{D}^{(2)} = \mathcal{D}^{(0)}\mathcal{D}^{(1)}\mathcal{D}^{(2)} \circ \mathcal{E}. \tag{C3}$$

The measurements comprising ${\rm Tr}_{(1,2)}\, M$ are of stabilizer operators $K^{(1)}_{\boldsymbol{0},\boldsymbol{b}}$, $K^{(2)}_{\boldsymbol{0},\boldsymbol{b}}$ on the states $\rho^{(1)}$, $\rho^{(2)}$, respectively. They are performed via one-qubit measurements and classical post-processing. $K^{(1)}_{\boldsymbol{0},\boldsymbol{b}}$, $K^{(2)}_{\boldsymbol{0},\boldsymbol{b}}$ commute with the Kraus operators in Eq. (22), such that

$$\begin{aligned} \underset{(1,2)}{\rm Tr}\, M \circ \mathcal{D}^{(0)}\mathcal{D}^{(1)}\mathcal{D}^{(2)} &= \underset{(1,2)}{\rm Tr}\, \mathcal{D}^{(0)}\mathcal{D}^{(1)}\mathcal{D}^{(2)} \circ M \\ &= \mathcal{D}^{(0)} \circ \underset{(1,2)}{\rm Tr}\, M. \end{aligned} \tag{C4}$$

Eqs. (C2), (C3), and (C4) yield Eq. (23)

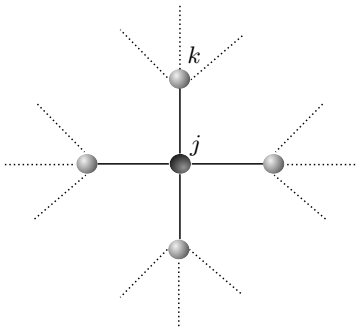

FIG. 5: Creation of a degree ($d = 4$) bicolorable graph state. The figure will have the same local structure for other degrees and topologies as long as its edges are d colorable and its vertices are bicolorable

APPENDIX D: CREATION OF A BI-COLORABLE GRAPH STATE

Here we discuss the noise structure of a bicolorable graph that is created using noisy CPHASE gates. The noisy gates are modeled as the ideal gate followed by two-qubit depolarizing noise as defined in Eq. (15). The graph state is created by performing CPHASE gates between qubits in the $|+\rangle$ state. The noise structure of the final state depends on the temporal ordering of these gates. If we assume that the underlying graph has constant degree d and that its edges are d colorable, then the N-qubit graph state can be created in d time steps with Nd CPHASE gates. At each time step all the gates corresponding to edges of a particular color are performed. Thus, at every time step $t \in \{1, \ldots, d\}$, each qubit is affected by an error channel of the form of Eq. (15).

We are interested in the value of $\langle K_j \rangle$, so we focus on the neighborhood of qubits j in the larger graph. Since the graph is bicolorable, it contains no three cycles and one can draw a diagram of the form of Fig. 5. The gates are represented by both solid as well as dashed lines. The noise channels corresponding to the solid lines each contribute an effective error ${\rm T}_{\rm eff}$ as defined in Eq. (16) to qubit j. Now consider the qubit k which is a neighbor of the central qubit j. Each dashed line also contributes an effective error ${\rm T}_{\rm eff}$ to qubit j, but only if the CPHASE gate corresponding to the solid line between k and j was performed in a previous timestep. This is because Z_k errors commute with K_j and X_k errors would be propagated by the CPHASE to $X_k Z_j$ errors,

which also commute with K_j. Thus there are a total of $\frac{d(d-1)}{2} + d = \frac{d(d+1)}{2}$ noise channels affecting the qubit j. This gives

$$\langle K_j \rangle = (1 - p_2)^{\frac{d(d+1)}{2}}. \tag{D1}$$

[1] A. K. Ekert, Phys. Rev. Lett. **67**, 661 (1991).
[2] K. Chen and H.-K. Lo (2004), arXiv:quant-ph/0404133.
[3] D. Gottesman and I. L. Chuang, Nature (London) **402**, 390 (1999).
[4] R. Raussendorf and H. J. Briegel, Phys. Rev. Lett. **86**, 5188 (2001), eprint, arXiv:quant-ph/0010033.
[5] C. H. Bennett, G. Brassard, S. Popescu, B. Schumacher, J. A. Smolin, and W. K. Wootters, Phys. Rev. Lett. **76**, 722 (1996).
[6] C. H. Bennett, D. P. DiVincenzo, J. A. Smolin, and W. K. Wootters, Phys. Rev. A **54**, 3824 (1996).
[7] H.-K. Lo and H. F. Chau, Science **283**, 2050 (1999).
[8] W. Dür, H. Aschauer, and H.-J. Briegel, Physical Review Letters **91**, 107903 (2003).
[9] H. Aschauer, W. Dür, and H.-J. Briegel, Phys. Rev. A **71**, 012319 (2005).
[10] E. Hostens, J. Dehaene, and B. D. Moor (2005), arXiv:quant-ph/0510096.
[11] A. Miyake and H. J. Briegel, Phys. Rev. Lett. **95**, 220501 (2005).
[12] W. Dür, H.-J. Briegel, J. I. Cirac, and P. Zoller, Phys. Rev. A **59**, 169 (1999).
[13] S. Bravyi (2005), arXiv:quant-ph/0511178.
[14] H. Aschauer and H. J. Briegel, Phys. Rev. Lett. **88**, 047902 (2002).
[15] B. M. Terhal and G. Burkard, Phys. Rev. A **71**, 012336 (2005).
[16] P. Aliferis, D. Gottesman, and J. Preskill, Quant. Inf. Comput. **6**, 097 (2006).
[17] R. Klesse and S. Frank, Phys. Rev. Lett. **95**, 230503 (2005).
[18] W. Dür, M. Hein, J. I. Cirac, and H.-J. Briegel, Phys. Rev. A **72**, 052326 (2005).
[19] C. Kruszynska, S. Anders, W. Dür, and H. J. Briegel (2005), arXiv:quant-ph/0512218.
[20] M. A. Nielsen and C. M. Dawson, Phys. Rev. A **71**, 042323 (2005), arXiv:quant-ph/0405134.
[21] C. M. Dawson, H. L. Haselgrove, and M. A. Nielsen, Phys. Rev. Lett. **96**, 020501 (2006), arXiv:quant-ph/0509060.
[22] P. Aliferis and D. W. Leung, Phys. Rev. A **73**, 032308 (2006).
[23] M. Varnava, D. E. Browne, and T. Rudolph (2005), arXiv:quant-ph/0507036.
[24] R. Raussendorf, J. Harrington, and K. Goyal (2005), arXiv:quant-ph/0510135.
[25] D. Smith and M. Latham, *The Geometry of Rene Descartes with a facsimile of the first edition* (Dover Publications, New York, 1954).
[26] For hashing, all N copies are included from the beginning. Each qubit of the state copies which are later measured is acted upon by a large number of noisy CNOT-gates. The error-correction procedure is applied only after the CNOTs have acted, such that their errors accumulate. Thus in the large N limit no matter how small the gate noise, the output state will be severely affected and the protocol will fail.